

COUNTING TREES USING SYMMETRIES

OLIVIER BERNARDI AND ALEJANDRO H. MORALES

ABSTRACT. We prove a new formula for the generating function of multitype Cayley trees counted according to their degree distribution. Using this formula we recover and extend several enumerative results about trees. In particular, we extend some results by Knuth and by Bousquet-Mélou and Chapuy about embedded trees. We also give a new proof of the multivariate Lagrange inversion formula. Our strategy for counting trees is to exploit symmetries of refined enumeration formulas: proving these symmetries is easy, and once the symmetries are proved the formulas follow effortlessly. We also adapt this strategy to recover an enumeration formula of Goulden and Jackson for cacti counted according to their degree distribution.

1. INTRODUCTION

The enumeration of trees is a very classical subject. For instance, there is a well-known formula for the number of *unitype Cayley trees*. Recall that a unitype Cayley tree with n vertices is a connected acyclic graph with vertex set $[n] = \{1, \dots, n\}$. There are n^{n-2} such trees, and there is a very simple formula for the generating function of Cayley trees counted according to their degree distribution. Namely,

$$(1) \quad \sum_{\substack{T \text{ Cayley tree} \\ \text{with vertex set } [n]}} x_1^{\deg(1)} x_2^{\deg(2)} \cdots x_n^{\deg(n)} = x_1 x_2 \cdots x_n (x_1 + x_2 + \cdots + x_n)^{n-2},$$

where $\deg(i)$ is the degree of vertex i .

In this paper we consider *multitype Cayley trees*, that is, trees in which vertices have both a *type* and a *label*. We obtain a formula extending (1) from the unitype setting to the multitype setting (Theorem 2). More precisely, our formula gives the generating function of rooted multitype Cayley trees counted according to the number of children of each type of each vertex. Our formula is surprisingly simple, and from it we derive many enumerative corollaries in Section 3. In particular, we recover and extend the results of Knuth [13], and the recent results of Bousquet-Mélou and Chapuy [4] about “embedded trees”. We also obtain a short proof of the multivariate Lagrange inversion formula [7] in Section 4. Our strategy for counting trees is to exploit symmetries of refined enumeration formulas, and we also use this strategy in order to recover a formula of Goulden and Jackson for counting cacti according to their degree distribution in Section 5. We mention lastly that because we count trees according to their vertex degrees, our results could equivalently be stated in terms of plane trees instead of Cayley trees (see Section 5 for a more detailed discussion). Also, our results can easily be extended in order to count rooted forests (see Corollary 3).

Date: April 8, 2013.

O.B. acknowledges support from NSF grant DMS-1068626, ANR A3, and ERC Explore-Maps.

In order to illustrate our approach for counting trees, we give a new proof of (1). There are already many beautiful proofs of this formula including Prüfer's code bijection [17], Joyal's endofunction approach [12], Pitman's double counting argument [16], the matrix-tree theorem [15, Chapter 5], and recursive approaches [18, Chapter 5.3]. Our method is different: we start by proving the “symmetries” in formula (1) and use them at our advantage in order to enumerate Cayley trees.

First observe that a Cayley tree with n vertices has $n - 1$ edges, hence the degrees of its vertices sum to $2n - 2$. Given a tuple of positive integers $\gamma = (\gamma_1, \dots, \gamma_n)$ summing to $2n - 2$, we denote by \mathcal{T}_γ the set of Cayley trees with n vertices such that vertex i has degree γ_i for all $i \in [n]$. We first claim that the cardinalities of the sets \mathcal{T}_γ are related to one another by simple factors:

Lemma 1. *Let $\gamma = (\gamma_1, \dots, \gamma_n)$ be tuple of positive integers summing to $2n - 2$. Let $i, j \in [n]$ and let $\gamma' = (\gamma'_1, \dots, \gamma'_n)$ be defined by $\gamma'_i = \gamma_i - 1$, $\gamma'_j = \gamma_j + 1$ and $\gamma'_k = \gamma_k$ for $k \neq i, j$. Then*

$$(\gamma_i - 1)|\mathcal{T}_\gamma| = (\gamma'_j - 1)|\mathcal{T}_{\gamma'}|.$$

Proof. The proof is illustrated in Figure 1. Let $\mathcal{T}_\gamma^{i,j}$ be the set of trees in \mathcal{T}_γ with a marked edge incident to vertex i not in the path between vertices i and j . Clearly $|\mathcal{T}_\gamma^{i,j}| = (\gamma_i - 1)|\mathcal{T}_\gamma|$. Moreover, there is an obvious bijection Φ between $\mathcal{T}_\gamma^{i,j}$ and $\mathcal{T}_{\gamma'}^{j,i}$: given a marked tree $T \in \mathcal{T}_\gamma^{i,j}$, the tree $\Phi(T) \in \mathcal{T}_{\gamma'}^{j,i}$ is obtained by ungluing the marked edge from vertex i , and gluing it to vertex j . \square

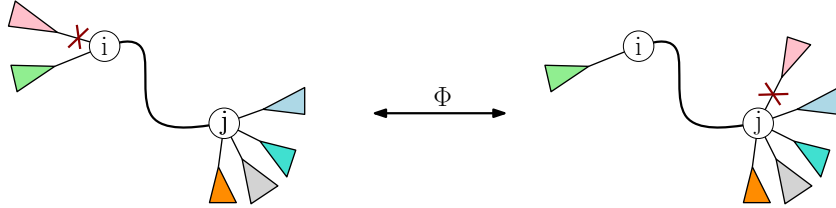


FIGURE 1. The bijection Φ between the sets $\mathcal{T}_\gamma^{i,j}$ and $\mathcal{T}_{\gamma'}^{j,i}$.

Using Lemma 1 repeatedly, we can express $|\mathcal{T}_\gamma|$ in terms of $|\mathcal{T}_\kappa|$, where $\kappa = (n - 1, 1, 1, \dots, 1)$. Indeed,

$$\begin{aligned} |\mathcal{T}_\gamma| &= \frac{\gamma_1(\gamma_1 + 1) \cdots (\gamma_1 + \gamma_2 - 2)}{(\gamma_2 - 1)!} |\mathcal{T}_{\gamma_1 + \gamma_2 - 1, 1, \gamma_3, \dots, \gamma_n}| \\ (2) \quad &= \frac{\gamma_1(\gamma_1 + 1) \cdots (\gamma_1 + \gamma_2 + \cdots + \gamma_n - n)}{(\gamma_2 - 1)! (\gamma_3 - 1)! \cdots (\gamma_n - 1)!} |\mathcal{T}_\kappa| \\ &= \binom{n - 2}{\gamma_1 - 1, \gamma_2 - 1, \dots, \gamma_n - 1} |\mathcal{T}_\kappa|. \end{aligned}$$

Moreover, $|\mathcal{T}_\kappa| = 1$ (only one “star tree”), hence $|\mathcal{T}_\gamma| = \binom{n - 2}{\gamma_1 - 1, \gamma_2 - 1, \dots, \gamma_n - 1}$. This implies (1) since this multinomial is the coefficient of $x_1^{\gamma_1} x_2^{\gamma_2} \cdots x_n^{\gamma_n}$ in $x_1 x_2 \cdots x_n (x_1 + x_2 + \cdots + x_n)^{n-2}$.

We will now use the above philosophy for tackling more advanced counting problems of tree-like structures.

2. THE GENERATING FUNCTION OF MULTITYPE CAYLEY TREES

Let $\mathbf{n} = (n_1, n_2, \dots, n_d)$ be a tuple of non-negative integers. A *multitype Cayley tree* of profile \mathbf{n} is a tree (i.e., acyclic connected graph) with vertex set

$$V_{\mathbf{n}} = \{(t, i), t \in [d], i \in [n_t]\}.$$

The vertex $(t, i) \in V_{\mathbf{n}}$ is said to have *type* t and *label* i . A multitype Cayley tree is represented in Figure 2.

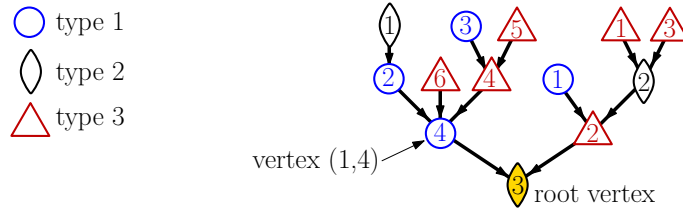


FIGURE 2. A multitype Cayley tree of profile $\mathbf{n} = (4, 3, 6)$ rooted on a vertex of type 2. The shape of the vertices indicate their type, while the numbers indicate their label. For the vertex $(1, 4)$ the indegrees are $\text{in}_1(1, 4) = 1$, $\text{in}_2(1, 4) = 0$, and $\text{in}_3(1, 4) = 2$.

A multitype Cayley tree is said to be *rooted* if one of the vertices is distinguished as the *root vertex*; in this case the edges of the tree are oriented toward the root vertex. We denote by $\mathcal{T}_\rho(\mathbf{n})$ the set of rooted multitype Cayley trees of profile \mathbf{n} in which the root vertex has type ρ . Given a tree $T \in \mathcal{T}_\rho(\mathbf{n})$, denote by $\text{in}_s(t, i)$ the number of children of type s of the vertex (t, i) . The tuple of integers $(\text{in}_s(t, i))_{s, t \in [d], i \in [n_t]}$ is called the *indegree vector* of T . We now state our main result.

Theorem 2. *Let $\mathbf{n} = (n_1, n_2, \dots, n_d)$ be a tuple of positive integers, and let $\rho \in [d]$. The generating function of rooted multitype Cayley trees of profile \mathbf{n} with root vertex of type ρ counted according to their indegree vectors is*

$$(3) \quad \sum_{T \in \mathcal{T}_\rho(\mathbf{n})} \prod_{s, t \in [d], i \in [n_t]} x_{s, t, i}^{\text{in}_s(t, i)} = \prod_{s \in [d]} \left(\sum_{t \in [d], i \in [n_t]} x_{s, t, i} \right)^{n_s - 1} \times \Delta,$$

where

$$\Delta = \sum_{A \in \text{Cayley}(d, \rho)} \prod_{(s, t) \in A} \left(\sum_{i \in [n_t]} x_{s, t, i} \right).$$

where $\text{Cayley}(d, \rho)$ is the set of unitype Cayley trees A with vertex set $[d]$ rooted at vertex ρ and considered as oriented toward its root vertex, and the notation $(s, t) \in A$ means that the oriented edge (s, t) belongs to the oriented tree A .

Remark. The set $\text{Cayley}(d, \rho)$ has cardinality d^{d-2} . This is the set of spanning trees of the complete graph K_d , with root vertex ρ , hence one can express the sum Δ appearing in (3) as a determinant by using the matrix-tree theorem [18, Theorem 5.6.8]. More precisely, let L be the $d \times d$ matrix with entries $\ell_{s,t} = -\sum_{i \in [n_t]} x_{s,t,i}$ if $s \neq t$, and $\ell_{s,s} = -\sum_{t \in [d], t \neq s} \ell_{s,t}$ for all $s \in [d]$. Then Δ is the determinant of the matrix obtained by deleting the ρ th row and ρ th column of L .

The case $d = 1$ of Theorem 2 corresponds to the enumeration of unitype *rooted* Cayley trees (i.e., rooted spanning trees of the complete graph K_n) according to the indegree of vertices. Indeed, upon setting $d = 1$, $n_1 = n$ and $x_{1,1,i} = x_i$ for all $i \in [n]$ in (3) one gets the well-known formula

$$\sum_{\substack{T \text{ rooted Cayley tree} \\ \text{with vertex set } [n]}} x_1^{\text{in}(1)} x_2^{\text{in}(2)} \cdots x_n^{\text{in}(n)} = (x_1 + x_2 + \cdots + x_n)^{n-1},$$

which is easily seen to be equivalent to (1). In the case $d = 2$, Theorem 2 can be specialized to give the generating function of the spanning trees of the complete bipartite graph $K_{m,n}$ counted according to the indegree of vertices; see e.g. [18, Exercise 5.30]. Indeed, upon setting $d = 2$, $n_1 = m$, $n_2 = n$, and $x_{2,1,i} = x_i$, $x_{1,2,j} = y_j$, $x_{1,1,i} = x_{2,2,j} = 0$ for all $i \in [m], j \in [n]$ one gets:

$$\sum_{T \subset K_{m,n}} \prod_{i \in [m]} x_i^{\text{in}(1,i)} \prod_{j \in [n]} y_j^{\text{in}(2,j)} = (x_1 + x_2 + \cdots + x_m)^n (y_1 + y_2 + \cdots + y_n)^{m-1},$$

where the sum is over all the spanning trees of $K_{m,n}$ rooted on a vertex of type 1. Many more applications are discussed in Section 3.

Before proving Theorem 2, we mention a corollary about rooted forests. Recall that a *forest* is an acyclic graph (hence each component is a tree), and that a forest is said to be *rooted* if each connected component has a vertex distinguished as the *root vertex*. We denote by $\mathcal{F}(\mathbf{n})$ the set of *rooted multitype forests of profile* $\mathbf{n} = (n_1, \dots, n_d)$, that is, the set of rooted forests with vertex set $\{(t, i), t \in [d], i \in [n_t]\}$. For $F \in \mathcal{F}(\mathbf{n})$ we think of each connected component of F as being oriented toward its root vertex, and we denote by $\text{in}_s(t, i)$ the number of children of type s of the vertex (t, i) .

Corollary 3. *Let $\mathbf{n} = (n_1, \dots, n_d)$ be a tuple of positive integers, and let*

$$\mathcal{F}_{\mathbf{n}}(\mathbf{x}, \mathbf{z}) = \sum_{F \in \mathcal{F}(\mathbf{n})} \prod_{s \in [d]} z_s^{\text{root}_s(F)} \prod_{s, t \in [d], i \in [n_t]} x_{s,t,i}^{\text{in}_s(t,i)},$$

where $\text{root}_s(F)$ is the number of root vertices of type s in the rooted forest F . Then,

$$(4) \quad \mathcal{F}_{\mathbf{n}}(\mathbf{x}, \mathbf{z}) = \prod_{s \in [d]} \left(z_s + \sum_{t \in [d], i \in [n_t]} x_{s,t,i} \right)^{n_s-1} \times \Gamma,$$

where

$$\Gamma = \sum_{B \in \mathcal{F}_d} \left(\prod_{s \text{ root vertex of } B} z_s \right) \prod_{(s,t) \in B} \left(\sum_{i \in [n_t]} x_{s,t,i} \right),$$

where F_d is the set of rooted forests with vertex set $[d]$ (where each connected component is thought as oriented toward its root vertex) and the notation $(s, t) \in B$ means that the oriented edge (s, t) belongs to the oriented forest B .

Remark. The set $\text{root}_s(F)$ has cardinality $(d+1)^{d-1}$. Using the (forest version of the) matrix-tree theorem one can express the sum Γ appearing in (4) as a determinant. More precisely Γ is the determinant of the $d \times d$ matrix L with entries $\ell_{s,t} = -\sum_{i \in [n_t]} x_{s,t,i}$ if $s \neq t$, and $\ell_{s,s} = z_s - \sum_{t \in [d], t \neq s} \ell_{s,t}$ for all $s \in [d]$.

Proof of Corollary 3. There is bijection between the set $\mathcal{F}(\mathbf{n})$ of rooted forests and the set $\mathcal{T}_{d+1}(\mathbf{n}')$ of rooted trees where $\mathbf{n}' = (n_1, \dots, n_d, 1)$. Indeed, given a forest $F \in \mathcal{F}(\mathbf{n})$ one gets a rooted tree $T \in \mathcal{T}_{d+1}(\mathbf{n}')$ by joining all the root vertices of F to a new vertex of type $d+1$ which becomes the root vertex of T . Thus,

$$\mathcal{F}_{\mathbf{n}}(\mathbf{x}, \mathbf{z}) = \sum_{T \in \mathcal{T}_{d+1}(\mathbf{n}')} \prod_{s \in [d]} z_s^{\text{in}_s(d+1,1)} \prod_{s,t \in [d], i \in [n_t]} x_{s,t,i}^{\text{in}_s(t,i)}.$$

Hence we can obtain the desired expression for $\mathcal{F}_{\mathbf{n}}(\mathbf{x}, \mathbf{z})$ by using (3) for the profile $\mathbf{n}' = (n_1, \dots, n_d, 1)$ and setting $x_{s,d+1,1} = z_s$. This immediately gives (4) because, via the bijection between F_d and $\text{Cayley}(d+1, d+1)$, one has

$$\Gamma = \sum_{A \in \text{Cayley}(d+1, d+1)} \left(\prod_{(s,d+1) \in A} z_s \right) \prod_{(s,t) \in A, t \in [d]} \left(\sum_{i \in [n_t]} x_{s,t,i} \right).$$

□

The rest of this section is devoted to the proof of Theorem 2. We start with the analogue of Lemma 1. Given a tuple $\gamma = (\gamma_{s,t,i})_{s,t \in [d], i \in [n_t]}$ of non-negative integers, we denote by $\mathcal{T}_{\rho, \gamma}$ the set of trees $T \in \mathcal{T}_{\rho}(\mathbf{n})$ having indegree vector γ , that is, satisfying $\text{in}_s(t, i) = \gamma_{s,t,i}$ for all $s, t \in [d], i \in [n_t]$. Observe that this set is empty unless,

$$n_s = \delta_{s,\rho} + \sum_{t \in [d]} \sum_{i \in [n_t]} \gamma_{s,t,i},$$

for all $s \in [d]$, where $\delta_{s,\rho}$ denotes the Kronecker delta.

Lemma 4. Let $\mathbf{n} = (n_1, n_2, \dots, n_d)$ be a tuple of positive integers and let $\gamma = (\gamma_{s,t,i})_{s,t \in [d], i \in [n_t]}$ be a tuple of non-negative integers. Let $s, t \in [d], i \neq j \in [n_t]$, and let $\gamma' = (\gamma'_{s,t,i})_{s,t \in [d], i \in [n_t]}$ be defined by $\gamma'_{s,t,i} = \gamma_{s,t,i} - 1$, $\gamma'_{s,t,j} = \gamma_{s,t,j} + 1$ and $\gamma'_{a,b,c} = \gamma_{a,b,c}$ for $(a, b, c) \notin \{(s, t, i), (s, t, j)\}$. Then

$$\gamma_{s,t,i} |\mathcal{T}_{\rho, \gamma}| = \gamma'_{s,t,j} |\mathcal{T}_{\rho, \gamma'}|.$$

Proof. Let $\mathcal{T}_{\rho, \gamma}^{s,t,i}$ be the set of trees in $\mathcal{T}_{\rho, \gamma}$ with a marked edge joining the vertex (t, i) to one of its children of type s . Clearly, $|\mathcal{T}_{\rho, \gamma}^{s,t,i}| = \gamma_{s,t,i} |\mathcal{T}_{\rho, \gamma}|$, hence it suffices to exhibit a bijection between $\mathcal{T}_{\rho, \gamma}^{s,t,i}$ and $\mathcal{T}_{\rho, \gamma'}^{s,t,j}$. We first partition $\mathcal{T}_{\rho, \gamma}^{s,t,i}$ into two sets $\widehat{\mathcal{T}}_{\rho, \gamma}^{s,t,i}$ and $\widetilde{\mathcal{T}}_{\rho, \gamma}^{s,t,i}$ defined as follows: a tree $T \in \mathcal{T}_{\rho, \gamma}^{s,t,i}$ is in $\widehat{\mathcal{T}}_{\rho, \gamma}^{s,t,i}$ if the marked edge e of T is not on the path of T between (t, i) and (t, j) (equivalently, e is not on the path from (t, j) to the root vertex), and the tree T is in $\widetilde{\mathcal{T}}_{\rho, \gamma}^{s,t,i}$ otherwise. We now describe a bijection

$\widehat{\Phi}_{s,t,i,j}$ between $\widehat{\mathcal{T}}_{\rho,\gamma}^{s,t,i}$ and $\widehat{\mathcal{T}}_{\rho,\gamma'}^{s,t,j}$, and a bijection $\widetilde{\Phi}_{s,t,i,j}$ between $\widetilde{\mathcal{T}}_{\rho,\gamma}^{s,t,i}$ and $\widetilde{\mathcal{T}}_{\rho,\gamma'}^{s,t,j}$. These bijections are represented in Figure 3.

Given a tree T in $\widehat{\mathcal{T}}_{\rho,\gamma}^{s,t,i}$, the tree $\widehat{\Phi}_{s,t,i,j}(T)$ is obtained by ungluing the marked edge from vertex (t,i) and gluing it to vertex (t,j) . It is clear that $\widehat{\Phi}_{s,t,i,j} = \widehat{\Phi}_{s,t,j,i}^{-1}$, hence $\widehat{\Phi}_{s,t,i,j}$ is a bijection between the sets $\widehat{\mathcal{T}}_{\rho,\gamma}^{s,t,i}$ and $\widehat{\mathcal{T}}_{\rho,\gamma'}^{s,t,j}$.

Given a tree T in $\widetilde{\mathcal{T}}_{\rho,\gamma}^{s,t,i}$, the tree $\widetilde{\Phi}_{s,t,i,j}(T)$ is obtained by ungluing all the unmarked edges oriented toward the vertex (t,i) and gluing them to (t,j) , ungluing all the edges originally oriented toward the vertex (t,j) and gluing them to (t,i) , and finally relabeling the vertex (t,i) as (t,j) and vice-versa; see Figure 3. It is clear that $\widetilde{\Phi}_{s,t,i,j} = \widetilde{\Phi}_{s,t,j,i}^{-1}$, hence $\widetilde{\Phi}_{s,t,i,j}$ is a bijection between the sets $\widetilde{\mathcal{T}}_{\rho,\gamma}^{s,t,i}$ and $\widetilde{\mathcal{T}}_{\rho,\gamma'}^{s,t,j}$. \square

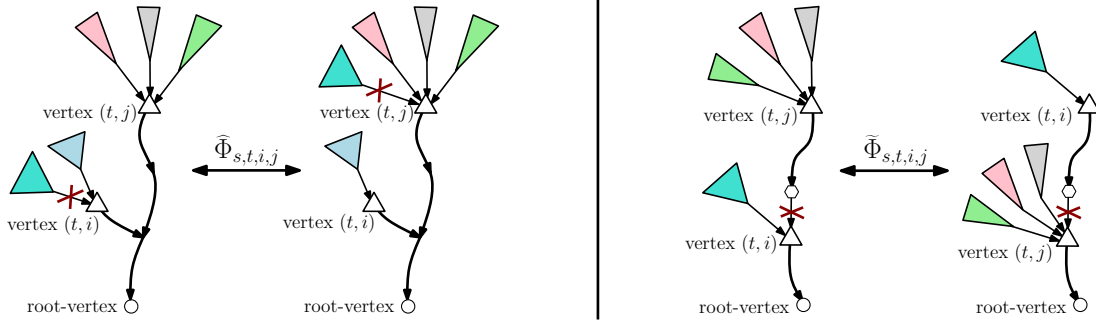


FIGURE 3. The bijections $\widehat{\Phi}_{s,t,i,j}$ (left) and $\widetilde{\Phi}_{s,t,i,j}$ (right).

A multitype Cayley tree is called a *star tree* if all the vertices not labeled 1 are leaves, that is, if $\text{in}_s(t,i) = 0$ for all $s, t \in [d]$ and all $i \neq 1$. The following Lemma shows that the problem of enumerating multitype Cayley trees reduces to the problem of enumerating star trees.

Lemma 5. *Let $\mathbf{n} = (n_1, n_2, \dots, n_d)$ be a tuple of positive integers. Let $\gamma = (\gamma_{s,t,i})_{s,t \in [d], i \in [n_t]}$ be a tuple of non-negative integers and let $\gamma^* = (\gamma_{s,t,i}^*)_{s,t \in [d], i \in [n_t]}$ be defined for all $s, t \in [d]$ by $\gamma_{s,t,1}^* = \sum_{i \in [n_t]} \gamma_{s,t,i}$ and $\gamma_{s,t,i}^* = 0$ for all $i \neq 1$. The number of trees of indegree vector γ is*

$$(5) \quad |\mathcal{T}_{\rho,\gamma}| = |\mathcal{T}_{\rho,\gamma^*}| \times \prod_{s,t \in [d]} \binom{\gamma_{s,t,1}^*}{\gamma_{s,t,1}, \dots, \gamma_{s,t,n_t}}.$$

Equivalently, in terms of generating functions,

$$(6) \quad \sum_{T \in \mathcal{T}_{\rho}(\mathbf{n})} \prod_{s,t \in [d], i \in [n_t]} x_{s,t,i}^{\text{in}_s(t,i)} = \sum_{T \text{ star tree in } \mathcal{T}_{\rho}(\mathbf{n})} \prod_{s,t \in [d]} \left(\sum_{i \in [n_t]} x_{s,t,i} \right)^{\text{in}_s(t,1)}.$$

Proof. Equation (5) is readily obtained by applying Lemma 4 repeatedly (in the spirit of what was done in Equation (2)). This implies (6) since extracting the coefficient of

$\prod_{s,t \in [d], i \in [n_t]} x_{s,t,i}^{\text{in}_s(t,i)}$ in the left-hand side of (6) gives $\mathcal{T}_{\rho,\gamma}$, while extracting this coefficient in the right-hand side gives

$$\sum_{T \in \mathcal{T}_{\rho,\gamma^*}} \prod_{s,t \in [d]} \binom{\gamma_{s,t,1}^*}{\gamma_{s,t,1}, \dots, \gamma_{s,t,n_t}}.$$

□

The following lemma completes the proof of Theorem 2.

Lemma 6. *Let $\mathbf{n} = (n_1, n_2, \dots, n_d)$ be a tuple of positive integers. The generating function of star trees of profile \mathbf{n} is given by*

$$(7) \quad \sum_{T \text{ star tree in } \mathcal{T}_{\rho}(\mathbf{n})} \prod_{s,t \in [d]} y_{s,t}^{\text{in}_s(t,1)} = \prod_{s \in [d]} \left(\sum_{t \in [d]} y_{s,t} \right)^{n_s-1} \times \sum_{A \in \text{Cayley}(d,\rho)} \prod_{(s,t) \in A} y_{s,t}.$$

Proof. Let T be a star tree in $\mathcal{T}_{\rho}(\mathbf{n})$. Since the vertices of T with labels distinct from 1 are leaves, removing these vertices gives a tree A with vertex set $\{(t, 1) \mid t \in [d]\}$ and root vertex $(\rho, 1)$. We call A the *core* of the star tree T and observe that the cores identify with the Cayley trees on $[d]$. Now, any star tree in $\mathcal{T}_{\rho}(\mathbf{n})$ is obtained by choosing a core A , and then adding the leaves (s, i) for $s \in [d]$ and $i \in [n_s] - \{1\}$. These leaves can be glued to any of the vertices $(1, 1), (2, 1), \dots, (d, 1)$ of A , and gluing a leaf to $(t, 1)$ increases $\text{in}_s(t, 1)$ by 1. This gives (7), where $\sum_{A \in \text{Cayley}(d,\rho)} \prod_{(s,t) \in A} y_{s,t}$ accounts

for choosing the core A , and $\prod_{s \in [d]} \left(\sum_{t \in [d]} y_{s,t} \right)^{n_s-1}$ accounts for adding the leaves. □

By Lemma 5, the generating function of multitype Cayley trees is obtained by substituting $y_{s,t}$ by $\sum_{i \in [n_t]} x_{s,t,i}$ in (7). This completes the proof of Theorem 2.

3. SOME APPLICATIONS

In this section we highlight a few consequences of Theorem 2. Throughout the section, $\mathbf{n} = (n_1, \dots, n_d)$ is a fixed tuple of positive integers, and ρ is in $[d]$. We say that an edge of a tree $T \in \mathcal{T}_{\rho}(\mathbf{n})$ has *type* (s, t) if it goes from a vertex of type s to a vertex of type t . Here is a summary of the results obtained below:

- In Proposition 7 we recover a result of Bousquet, Chauve, Labelle and Leroux [3, Proposition 2] by enumerating trees in $\mathcal{T}_{\rho}(\mathbf{n})$ with prescribed number $m_{s,t}$ of edges of type (s, t) .
- In Proposition 8 we recover a result by Knuth [13] by counting trees in $\mathcal{T}_{\rho}(\mathbf{n})$ such that the edges all have a type in a prescribed set $D \subset [d]^2$. These are called *D-embedded trees*.
- In Proposition 9 we enumerate *injective trees* in $\mathcal{T}_{\rho}(\mathbf{n})$, that is, trees such that every vertex has at most one child of each type. Our formula generalizes a result by Bousquet-Mélou and Chapuy [4, Theorem 6].
- In Proposition 10, we enumerate injective *D-embedded trees* in $\mathcal{T}_{\rho}(\mathbf{n})$.

- In Proposition 12 we enumerate trees in $\mathcal{T}_\rho(\mathbf{n})$ with prescribed number $N_{t,\mathbf{c}}$ of vertices of type t having *indegree vector* $\mathbf{c} = (c_1, \dots, c_d)$ (that is, having c_s children of type s). Our formula generalizes [4, Theorem 7].
- In Proposition 13 we enumerate trees in $\mathcal{T}_\rho(\mathbf{n})$ with prescribed number $N_{t,u,\mathbf{c}}$ of vertices of type t having a parent of type u and indegree vector \mathbf{c} . Our formula answers a question raised in [4].

We mention that a version of these results could be obtained for rooted forests using Corollary 3.

Throughout the section, we denote by $\text{Cayley}(d, \rho)$ the set of unitype Cayley trees with vertex set $[d]$ rooted at vertex ρ and considered as oriented toward their root vertex, and for a tree $A \in \text{Cayley}(d, \rho)$ the notation $(s, t) \in A$ indicates that the oriented edge (s, t) belongs to the oriented tree A . We first count multitype Cayley trees according to the number of edges of each type. For a tuple $\mathbf{m} = (m_{s,t})_{s,t \in [d]}$ of non-negative integers, we denote by $\mathcal{T}_\rho(\mathbf{m}, \mathbf{n})$ the set of trees in $\mathcal{T}_\rho(\mathbf{n})$ having $m_{s,t}$ edges of type (s, t) . In order for $\mathcal{T}_\rho(\mathbf{m}, \mathbf{n})$ to be non-empty we must impose

$$(8) \quad \text{for all } s \in [d], \quad n_s = \delta_{s,\rho} + \sum_{t \in [d]} m_{s,t},$$

where $\delta_{s,\rho}$ denotes the Kronecker delta. The following result was obtained by Bousquet, Chauve, Labelle and Leroux in [3, Proposition 2].

Proposition 7 ([3]). *Let $\mathbf{m} = (m_{s,t})_{s,t \in [d]}$ be a tuple of non-negative integers. The number of trees in $\mathcal{T}_\rho(\mathbf{n})$ with $m_{s,t}$ edges of type (s, t) for all $s, t \in [d]$ is*

$$(9) \quad |\mathcal{T}_\rho(\mathbf{m}, \mathbf{n})| = \left(\prod_{s,t \in [d]} n_t^{m_{s,t}} \right) \left(\frac{\prod_{s \in [d]} (n_s - 1)!}{\prod_{s,t \in [d]} m_{s,t}!} \right) \times \sum_{A \in \text{Cayley}(d, \rho)} \prod_{(s,t) \in A} m_{s,t}$$

if (8) holds and $|\mathcal{T}_\rho(\mathbf{m}, \mathbf{n})| = 0$ otherwise.

Remark. By the matrix-tree theorem, the sum over Cayley trees in (9) can be expressed as a determinant. More precisely, this sum is the determinant of the matrix obtained by deleting the ρ th row and ρ th column of the $d \times d$ matrix L having entries $\ell_{s,s} = \sum_{t \in [d], t \neq s} m_{s,t}$ for all $s \in [d]$, and $\ell_{s,t} = -m_{s,t}$ for all $t \neq s \in [d]$. Similar determinantal expressions exist for the formulas given in Propositions 8 to 13 and are omitted.

Proof. Setting $x_{s,t,i} = x_{s,t}$ for all i in (3) gives

$$\sum_{T \in \mathcal{T}_\rho(\mathbf{n})} \prod_{s,t \in [d]} x_{s,t}^{\#\text{edges of type } (s,t)} = \prod_{s \in [d]} \left(\sum_{t \in [d]} n_t x_{s,t} \right)^{n_s - 1} \times \sum_{A \in \text{Cayley}(d, \rho)} \prod_{(s,t) \in A} n_t x_{s,t},$$

and extracting the coefficient of $\prod_{s,t \in [d]} x_{s,t}^{m_{s,t}}$ in this equation gives

$$|\mathcal{T}_\rho(\mathbf{m}, \mathbf{n})| = \left(\prod_{s,t \in [d]} n_t^{m_{s,t}} \right) \sum_{A \in \text{Cayley}(d, \rho)} \prod_{s \in [d]} \left[\prod_{t \in [d]} x_{s,t}^{m_{s,t}} \right] \left(\sum_{t \in [d]} n_t x_{s,t} \right)^{n_s - 1} \prod_{t, (s,t) \in A} x_{s,t},$$

where the bracket notation means extraction of coefficients. Moreover, for any $s \in [d]$ there is at most one t with $(s, t) \in A$ (this is the parent of s in A) so that

$$\left[\prod_{t \in [d]} x_{s,t}^{m_{s,t}} \right] \left(\sum_{t \in [d]} n_t x_{s,t} \right)^{n_s-1} \prod_{t, (s,t) \in A} x_{s,t} = \frac{\prod_{s \in [d]} (n_s - 1)!}{\prod_{s,t \in [d]} m_{s,t}!} \prod_{t, (s,t) \in A} m_{s,t},$$

provided (8) holds. This gives (9). \square

Next, we count *embedded trees*. Given a subset D of $[d]^2$, we say that a multitype Cayley tree T is *embedded* in D if the type of every edge of T belongs to D . The reason for this terminology is that D can be seen as a digraph with vertex set $[d]$ and if the vertices of T are sent to the vertices of D corresponding to their type (i.e., $(t, i) \mapsto t$), then the edges of T are sent to edges of D . Let $\mathcal{T}_{\rho,D}(\mathbf{n})$ be the set of trees in $\mathcal{T}_{\rho}(\mathbf{n})$ embedded in D . The result proved by Knuth in [13] (using a variation of Joyal's endofunction technique [12]) is the following enumeration of $\mathcal{T}_{\rho,D}(\mathbf{n})$.

Proposition 8 ([13]). *Let $D \subseteq [d]^2$. The number of trees in $\mathcal{T}_{\rho}(\mathbf{n})$ embedded in D is*

$$|\mathcal{T}_{\rho,D}(\mathbf{n})| = \prod_{s \in [d]} \left(\sum_{(s,t) \in D} n_t \right)^{n_s-1} \times \sum_{A \in \text{Cayley}(d,\rho), A \subseteq D} \prod_{(s,t) \in A} n_t.$$

Proof. Setting $x_{s,t,i} = 1$ if $(s, t) \in D$ and $x_{s,t,i} = 0$ otherwise in (3) directly gives the result. \square

Next, we count *injective trees* and *embedded injective trees*. A multitype Cayley tree T is said to be *injective* if every vertex has at most one child of each type. Injective trees were introduced by Bousquet-Mélou and Chapuy in [4] in order to study the so-called *vertical profile* of trees. Let $\mathcal{T}_{\rho}^{\text{inj}}(\mathbf{m}, \mathbf{n})$ (resp. $\mathcal{T}_{\rho,D}^{\text{inj}}(\mathbf{n})$) be the subset of injective trees in $\mathcal{T}_{\rho}(\mathbf{m}, \mathbf{n})$ (resp. $\mathcal{T}_{\rho,D}(\mathbf{n})$). The following result generalizes [4, Theorem 6].

Proposition 9. *Let $\mathbf{m} = (m_{s,t})_{s,t \in [d]}$ be a tuple of non-negative integers. The number of injective trees in $\mathcal{T}_{\rho}(\mathbf{n})$ with $m_{s,t}$ edges of type (s, t) for all $s, t \in [d]$ is*

$$|\mathcal{T}_{\rho}^{\text{inj}}(\mathbf{m}, \mathbf{n})| = \prod_{s,t \in [d]} \binom{n_t}{m_{s,t}} \prod_{s \in [d]} (n_s - 1)! \times \sum_{A \in \text{Cayley}(d,\rho)} \prod_{(s,t) \in A} m_{s,t},$$

if (8) holds and $|\mathcal{T}_{\rho}^{\text{inj}}(\mathbf{m}, \mathbf{n})| = 0$ otherwise.

Proof. In order to choose a tree in $\mathcal{T}_{\rho}^{\text{inj}}(\mathbf{m}, \mathbf{n})$, one must first choose for all $s, t \in [d]$, the set $M_{s,t} \subseteq [n_t]$ of labels of the $m_{s,t}$ vertices of type t having a child of type s . There are $\prod_{s,t \in [d]} \binom{n_t}{m_{s,t}}$ such choices. Moreover, the number of trees in $\mathcal{T}_{\rho}^{\text{inj}}(\mathbf{m}, \mathbf{n})$ corresponding to a given choice $(M_{s,t})_{s,t \in [d]}$ of labels is the coefficient of $\prod_{s,t \in [d], i \in M_{s,t}} x_{s,t,i}$ in (3) which is easily seen to be $\prod_{s \in [d]} (n_s - 1)! \times \sum_{A \in \text{Cayley}(d,\rho)} \prod_{(s,t) \in A} m_{s,t}$, provided (8) holds. \square

The following result generalizes [4, Theorem 4].

Proposition 10. *Let $D \subseteq [d]^2$. The number of injective trees in $\mathcal{T}_\rho(\mathbf{n})$ embedded in D is*

$$|\mathcal{T}_{\rho,D}^{\text{inj}}(\mathbf{n})| = \prod_{s \in [d]} \binom{\delta_{s,\rho} - 1 + \sum_{t \in D(s)} n_t}{n_s - 1} (n_s - 1)! \times \sum_{A \subseteq D \text{ in Cayley}(d,\rho)} \prod_{(s,t) \in A} n_t.$$

where $D(s) = \{t \in [d] \mid (s,t) \in D\}$.

Proof. By (3), $|\mathcal{T}_{\rho,D}^{\text{inj}}(\mathbf{n})|$ is the number of square-free monomials in the expansion of

$$\prod_{s \in [d]} \left(\sum_{t \in D(s), i \in [n_t]} x_{s,t,i} \right)^{n_s - 1} \times \sum_{A \subseteq D \text{ in Cayley}(d,\rho)} \prod_{(s,t) \in A} \left(\sum_{i \in [n_t]} x_{s,t,i} \right).$$

Expanding the sum over A gives a sum of

$$\sum_{A \subseteq D \text{ in Cayley}(d,\rho)} \prod_{(s,t) \in A} n_t$$

square-free monomials of the form $\prod_{s \in [d] \setminus \{\rho\}} y_s$, with y_s in $X_s = \{x_{s,t,i}\}_{t \in D(s), i \in [n_t]}$. Moreover, there are clearly

$$\prod_{s \in [d]} \binom{\delta_{s,\rho} - 1 + \sum_{t \in D(s)} n_t}{n_s - 1} (n_s - 1)!$$

square-free monomials not containing the variable y_s for all $s \in [d] \setminus \{\rho\}$ in the expansion of

$$\prod_{s \in [d]} \left(\sum_{t \in D(s), i \in [n_t]} x_{s,t,i} \right)^{n_s - 1} = \prod_{s \in [d]} \left(\sum_{y \in X_s} y \right)^{n_s - 1}.$$

□

We will now count trees according to their *indegree vectors* and *complete degree vectors*. Recall that for a tuple $\gamma = (\gamma_{s,t,i})_{s,t \in [d], i \in [n_t]}$ of non-negative integers, $\mathcal{T}_{\rho,\gamma}$ denotes the set of trees in $\mathcal{T}_\rho(\mathbf{n})$ such that $\text{in}_s(t,i) = \gamma_{s,t,i}$ for all $s, t \in [d], i \in [n_t]$.

Proposition 11. *Let $\gamma = (\gamma_{s,t,i})_{s,t \in [d], i \in [n_t]}$ be a tuple of non-negative integers, and let $m_{s,t} = \sum_{i \in [n_t]} \gamma_{s,t,i}$. The number of trees with indegree vector γ (hence having $m_{s,t}$ edges of type (s,t)) is*

$$(10) \quad |\mathcal{T}_{\rho,\gamma}| = \frac{\prod_{t \in [d]} (n_t - 1)!}{\prod_{s,t \in [d], i \in [n_t]} \gamma_{s,t,i}!} \times \sum_{A \in \text{Cayley}(d,\rho)} \prod_{(s,t) \in A} m_{s,t},$$

if (8) holds, and $|\mathcal{T}_{\rho,\gamma}| = 0$ otherwise.

Proof. By extracting the coefficient of $\prod_{s,t \in [d], i \in [n_t]} x_{s,t,i}^{\gamma_{s,t,i}}$ in (3) we get

$$\begin{aligned} |\mathcal{T}_{\rho,\gamma}| &= \left[\prod_{s,t \in [d], i \in [n_t]} x_{s,t,i}^{\gamma_{s,t,i}} \right] \sum_{A \in \text{Cayley}(d,\rho)} \left(\prod_{(s,t) \in A} \sum_{i \in [n_t]} x_{s,t,i} \right) \prod_{s \in [d]} \left(\sum_{t \in [d], i \in [n_t]} x_{s,t,i} \right)^{n_s - 1} \\ (11) \quad &= \sum_{A \in \text{Cayley}(d,\rho)} \prod_{s \in [d]} C_{A,s} \end{aligned}$$

where

$$C_{A,s} = \left[\prod_{t \in [d], i \in [n_t]} x_{s,t,i}^{\gamma_{s,t,i}} \right] \left(\prod_{t, (s,t) \in A} \sum_{i \in [n_t]} x_{s,t,i} \right) \left(\sum_{t \in [d], i \in [n_t]} x_{s,t,i} \right)^{n_s-1}.$$

We now assume that (8) holds. Since there is no $t \in [d]$ such that $(\rho, t) \in A$,

$$C_{A,\rho} = \left[\prod_{t \in [d], i \in [n_t]} x_{\rho,t,i}^{\gamma_{\rho,t,i}} \right] \left(\sum_{t \in [d], i \in [n_t]} x_{\rho,t,i} \right)^{n_\rho-1} = \frac{(n_\rho-1)!}{\prod_{t \in [d], i \in [n_t]} \gamma_{\rho,t,i}}.$$

Similarly, by denoting s' the parent of a vertex $s \neq \rho$ in A , we get

$$C_{A,s} = \left[\prod_{t \in [d], i \in [n_t]} x_{s,t,i}^{\gamma_{s,t,i}} \right] \sum_{i \in [n_{s'}]} x_{s,s',i} \left(\sum_{t \in [d], i \in [n_t]} x_{s,t,i} \right)^{n_s-1} = \left(\sum_{i \in [n_{s'}]} \gamma_{s,s',i} \right) \frac{(n_s-1)!}{\prod_{t \in [d], i \in [n_t]} \gamma_{s,t,i}}.$$

Using $\sum_{i \in [n_{s'}]} \gamma_{s,s',i} = m_{s,s'}$, and plugging the expression of $C_{A,s}$ in (11) gives (10). \square

Another way of stating (10) is by fixing the number of vertices of each indegree type (but without fixing their labels). We say that a vertex of a tree $T \in \mathcal{T}_\rho(\mathbf{n})$ has *indegree type* $\mathbf{c} = (c_1, \dots, c_d)$ if it has c_s children of type s for all $s \in [d]$. For instance, the vertex (1, 4) in Figure 2 has indegree type $\mathbf{c} = (1, 0, 2)$. Let $\mathbf{C} = [n_1] \times \dots \times [n_d]$, and let $\mathbf{N} = (N_{t,\mathbf{c}})_{t \in [d], \mathbf{c} \in \mathbf{C}}$ be a tuple of non-negative integers. Let $\mathcal{T}_\rho^{\mathbf{N}}$ be the set of trees in $\mathcal{T}_\rho(\mathbf{n})$ having $N_{t,\mathbf{c}}$ vertices of type t with indegree type \mathbf{c} , for all $t \in [d], \mathbf{c} \in \mathbf{C}$. The following result generalizes [4, Theorem 7].

Proposition 12. *Let $\mathbf{N} = (N_{t,\mathbf{c}})_{t \in [d], \mathbf{c} \in \mathbf{C}}$ be a tuple of non-negative integers. Let $n_t = \sum_{\mathbf{c} \in \mathbf{C}} N_{t,\mathbf{c}}$ for all $t \in [d]$, let $m_{s,t} = \sum_{\mathbf{c} \in \mathbf{C}} c_s N_{t,\mathbf{c}}$, and let $N(k) = \sum_{\substack{s,t \in [d], \mathbf{c} \in \mathbf{C}, \\ \text{such that } c_s = k}} N_{t,\mathbf{c}}$.*

The number of trees in $\mathcal{T}_\rho(\mathbf{n})$ having $N_{t,\mathbf{c}}$ vertices of type t with indegree type \mathbf{c} is

$$|\mathcal{T}_\rho^{\mathbf{N}}| = \frac{\prod_{t \in [d]} n_t! (n_t - 1)!}{\prod_{t \in [d], \mathbf{c} \in \mathbf{C}} N_{t,\mathbf{c}}! \prod_{k \geq 0} k!^{N(k)}} \times \sum_{A \in \text{Cayley}(d, \rho)} \prod_{(s,t) \in A} m_{s,t},$$

if (8) holds and $|\mathcal{T}_\rho^{\mathbf{N}}| = 0$ otherwise.

Proof. By definition, $\mathcal{T}_\rho^{\mathbf{N}} = \bigcup_{\gamma \in \Gamma(\mathbf{N})} \mathcal{T}_{\rho,\gamma}$, where $\Gamma(\mathbf{N})$ is the set of tuples on non-negative integers $(\gamma_{s,t,i})_{s,t \in [d], i \in [n_t]}$ such that for all $t \in [d]$ and all $\mathbf{c} \in \mathbf{C}$, there are $N_{t,\mathbf{c}}$ integers $i \in [n_t]$ with $(\gamma_{1,t,i}, \dots, \gamma_{d,t,i}) = \mathbf{c}$. Since $|\Gamma(\mathbf{N})| = \frac{\prod_{t \in [d]} n_t!}{\prod_{t \in [d], \mathbf{c} \in \mathbf{C}} N_{t,\mathbf{c}}!}$, using (10) gives the result. \square

Remark. Propositions 7 to 12 give particularly nice counting formulas (which factorize completely) when there is a unique tree $A \in \text{Cayley}(d, \rho)$ contributing to the sum (the cases considered in [4] are all of this form). It is therefore interesting to understand when this favorable situation occurs. A quick investigation reveals that it occurs when

the type of the edges allowed to appear in the trees to be enumerated belong to a set $D \subseteq [d]^2$ which can be partitioned as $D = A \cup A'$, where A is the edge set of any tree in $\text{Cayley}(d, \rho)$ and A' is a set of edges of the form (s, t) where $t = s$ or t is a descendant of s in A (i.e., (t, s) is in A). For instance, if $A \in \text{Cayley}(d, \rho)$, and $m_{s,t} = 0$ unless (s, t) or (t, s) is in A , then Proposition 12 gives

$$|\mathcal{T}_\rho^{\mathbf{N}}| = \frac{\prod_{t \in [d]} n_t! (n_t - 1)!}{\prod_{t \in [d], c \in \mathbf{C}} N_{t,c}! \prod_{k \geq 0} k!^{N(k)}} \prod_{(s,t) \in A} m_{s,t}.$$

As this section's last application of Theorem 2, we count trees according to their complete degree type. We say that a vertex of a tree in $\mathcal{T}_\rho(\mathbf{n})$ has *complete degree type* (t, u, \mathbf{c}) if it has type t , indegree type \mathbf{c} and its parent has type u , with the convention that the fictitious “parent” of the root vertex has type $d + 1$. For instance, the vertex $(1, 4)$ in Figure 2 has complete degree type $(1, 2, (1, 0, 2))$, while the root vertex has complete degree type $(2, 4, (1, 0, 1))$. Let $\vec{\mathcal{T}}_\rho^{\mathbf{N}}$ be the set of multitype Cayley trees with a root vertex of type ρ having $N_{t,u,\mathbf{c}}$ vertices of complete degree type (t, u, \mathbf{c}) for all $t \in [d], u \in [d + 1], \mathbf{c} \in \mathbf{C}$.

Proposition 13. *Let $\mathbf{N} = (N_{t,u,\mathbf{c}})_{t \in [d], u \in [d+1], \mathbf{c} \in \mathbf{C}}$ be a tuple of non-negative integers. Let $m_{t,u} = \sum_{\mathbf{c} \in \mathbf{C}} N_{t,u,\mathbf{c}}$, let $m_{s,t,u} = \sum_{\mathbf{c} \in \mathbf{C}} c_s N_{t,u,\mathbf{c}}$, let $n_t = \sum_{u \in [d+1]} m_{t,u}$, and let*

$$N(k) = \sum_{\substack{s,t \in [d], u \in [d+1], \mathbf{c} \in \mathbf{C}, \\ \text{such that } c_s = k}} N_{t,u,\mathbf{c}}.$$

The number of trees in $\mathcal{T}_\rho(\mathbf{n})$ having $N_{t,u,\mathbf{c}}$ vertices of complete degree type (t, u, \mathbf{c}) is

$$(12) \quad |\vec{\mathcal{T}}_\rho^{\mathbf{N}}| = \frac{\prod_{t \in [d]} n_t! \prod_{s,t \in [d], m_{s,t} > 0} (m_{s,t} - 1)!}{\prod_{t,u \in [d], \mathbf{c} \in \mathbf{C}} N_{t,u,\mathbf{c}}! \prod_{k \geq 0} k!^{N(k)}} \times \sum_{A \subseteq G(\mathbf{N})} \prod_{((s,t),(t,u)) \in A} m_{s,t,u},$$

provided that $m_{s,d+1} = \delta_{s,\rho}$, and $m_{s,t} = \sum_{u \in [d+1]} m_{s,t,u}$ for all $s, t \in [d]$ ($|\vec{\mathcal{T}}_\rho^{\mathbf{N}}| = 0$ otherwise), where $G(\mathbf{N})$ is the graph with vertex set $V = \{(t, u) \in [d] \times [d + 1], m_{t,u} > 0\}$ and edge set $\{((s, t), (t, u)) \mid (s, t), (t, u) \in V\}$, and the sum is taken over the spanning trees A of $G(\mathbf{N})$ oriented toward the vertex $(\rho, d + 1)$.

As an illustration of Proposition 13, consider the case $d = 1$, $n_1 = n$ and $m_{1,1} = n - 1 > 0$. In this case, the graph $G(\mathbf{N})$ has two vertices $(1, 1)$ and $(1, 2)$ and two edges: one edge going from $(1, 1)$ to $(1, 2)$ and one edge going from $(1, 1)$ to $(1, 1)$. Hence the sum in (12) reduces to $m_{1,1,2}$ which is the specified degree of the root vertex. Equation (12) then gives the number $|\vec{\mathcal{T}}_\rho^{\mathbf{N}}|$ of Cayley trees with n vertices, with N_c non-root vertices of indegree c and a root vertex of degree ℓ as

$$|\vec{\mathcal{T}}_\rho^{\mathbf{N}}| = \frac{n!(n - 2)!}{\prod_{c \geq 0} N_c! c!^{N_c} (\ell - 1)!}.$$

Proof. The proof of Proposition 13 is based on the observation that counting trees according to complete degree types can be seen as a special case of counting trees according to indegree types. Indeed, let us define the *complete type* of a vertex v of a tree $T \in \mathcal{T}_\rho(\mathbf{n})$ as the pair $(t, u) \in [d] \times [d+1]$, where t is the type of v and u is the type of the parent of v , with the convention that $u = d+1$ if v is the root vertex. We can then directly apply Proposition 12 in order to count trees according to their complete degree type (upon observing that the number of vertices of complete type (s, t) is $m_{s,t}$, and the number of edges from a vertex of complete type (s, t) to a vertex of complete type (t, u) is $m_{s,t,u}$). We only need to remember that the vertices of our Cayley trees are still labeled within their “original type” and not within their complete type (that is, the labels of vertices of complete type (s, t) have a label in $[n_t]$ and not in $[m_{s,t}]$), which is accounted for by a factor $\frac{\prod_{t \in [d]} n_t!}{\prod_{s,t \in [d]} m_{s,t}!}$ in (12). \square

Observe that using the same techniques as for Proposition 13 one could enumerate multitype Cayley trees according to the number of vertices having given indegree type, given type, with parent of given type, and grandparent of given type, etc. We now investigate a case where the sum in (12) simplifies greatly. The following result answers a question raised in [4, Section 8.2].

Corollary 14. *Suppose, using the notation of Proposition 13, that $\rho = d$ and all the pairs $(s, t) \in [d] \times [d+1]$ such that $m_{s,t} > 0$ satisfy $t \leq s+1$. In this case,*

$$(13) \quad |\vec{\mathcal{T}}_d^{\mathbf{N}}| = \frac{\prod_{t \in [d]} n_t! \prod_{s,t \in [d], m_{s,t} > 0} (m_{s,t} - 1)!}{\prod_{t,u \in [d], c \in \mathbf{C}} N_{t,u,c}! \prod_{k \geq 0} k!^{N(k)}} \times \prod_{\substack{s,t \in [d], \text{ such that } \\ t < s \text{ and } m_{s,t} > 0}} m_{s,t} \\ \times \mu_1 \prod_{s=2}^d (m_{s-1,s,s+1} \mu_s + m_{s-1,s,s} m_{s,s,s+1}),$$

where $\mu_s = m_{s,s} - m_{s,s,s}$ if $m_{s,s} > 0$ and $\mu_s = 1$ if $m_{s,s} = 0$.

The result given in [4, Theorem 8] corresponds to (13) in the case where $m_{s,s} = 0$ for all s (in this case, $\mu_s = 1$ and $m_{s-1,s,s} = 0$). We now briefly explain how to derive Corollary 14 from Proposition 13. Let G' be the directed graph with vertex set $V' = \{(s, t) \in [d] \times [d+1] \mid t \leq s+1\}$ and edge set $E' = \{((s, t), (t, u)) \mid s, t \in [d], u \in [d+1]\}$. We denote $e_{s,t,u}$ the edge $((s, t), (t, u)) \in E'$. By hypothesis, the graph $G(\mathbf{N})$ is a subgraph of G' .

Claim 15. *Let A be a subset of E' , such that A contains exactly one edge going out of each vertex $(s, t) \in [d]^2$ (that is, an edge of the form $e_{s,t,u}$), but contains no loop (that is, no edge of the form $e_{s,s,s}$). Then A is a spanning tree of G' oriented toward the root vertex $(d, d+1)$ if and only if for all $s \in \{2, \dots, d\}$, A contains either the edge $e_{s-1,s,s+1}$ or both the edges $e_{s-1,s,s}$ and $e_{s,s,s+1}$.*

Proof. We consider the *lexicographic order* on the set V' of vertices (in this order $(s, t) \leq_{\text{lex}} (s', t')$ if and only if $s < s'$ or $s = s'$ and $t \leq t'$). Observe that the root vertex $(d, d+1)$ is the largest vertex in the lexicographic order and that A is a spanning

tree if and only if the unique directed path in A starting from any vertex $(s, t) \in [d]^2$ leads to the root vertex $(d, d+1)$. Suppose first that A is a spanning tree of G' oriented toward the root vertex $(d, d+1)$. Since $e_{s-1, s, s+1}$ and $e_{s-1, s, s}$ are the only edges in E' going from a vertex $(t, u) \leq_{\text{lex}} (s-1, s)$ to a vertex $(u, v) >_{\text{lex}} (s-1, s)$, we know that one of these edges is in A . Similarly, since $e_{s-1, s, s+1}$ and $e_{s, s, s+1}$ are the only edges in E' going from a vertex $(t, u) \leq_{\text{lex}} (s, s)$ to a vertex $(u, v) >_{\text{lex}} (s, s)$, we know that one of these edges is in A . Suppose conversely that for all $s \in \{2, \dots, d\}$, A contains either the edge $e_{s-1, s, s+1}$ or both the edges $e_{s-1, s, s}$ and $e_{s, s, s+1}$. We want to prove that for each $(s, t) \in [d]^2$ the directed path in A starting from (s, t) leads to the root vertex $(d, d+1)$. This property is obvious in the case $t = s+1$. Moreover in the case $t \leq s$ we observe that the vertices decrease strictly for the lexicographic order along the directed path starting at (s, t) until a vertex of the form $(s, s+1)$ is reached. This concludes the proof. \square

We now complete the proof of Corollary 14. Using Claim 15, one can decompose the sum in (12) according to the fact that the spanning tree A of $G(\mathbf{N})$ contains the edge $e_{s-1, s, s+1}$ or both the edges $e_{s-1, s, s}$ and $e_{s, s, s+1}$. We now claim that in the rightmost product of (13) the case $e_{s-1, s, s+1} \in A$ is accounted for by the term $m_{s-1, s, s+1} \mu_s$, while the case $\{e_{s-1, s, s}, e_{s, s, s+1}\} \subseteq A$ is accounted for by the term $m_{s-1, s, s} m_{s, s, s+1}$. We leave the details of this statement to the reader, and only mention that the term μ_s accounts for the choice of the edge of A going out of the vertex (s, s) (in the case $(s, s) \in V$), while the term $m_{s, t}$ in the first product accounts for the choice of the edge of A going out of the vertex (s, t) (in the case $(s, t) \in V$ with $t < s$).

4. MULTIVARIATE LAGRANGE INVERSION FORMULA

In this section we show that Theorem 2, via Proposition 11, implies the multivariate Lagrange inversion formula [8]. There are several versions of this formula which have been shown to be equivalent to each other; see [7] for a survey. Here we will derive a version due to Bender and Richmond [1].

We will consider power series in the variables x_1, \dots, x_d , and we denote $\mathbf{x} = (x_1, \dots, x_d)$. For a tuple of integers $\mathbf{n} = (n_1, \dots, n_d)$, we denote $\mathbf{x}^{\mathbf{n}} = x_1^{n_1} x_2^{n_2} \dots x_d^{n_d}$, and we denote $[\mathbf{x}^{\mathbf{n}}]f(\mathbf{x})$ for the coefficient of the monomial $\mathbf{x}^{\mathbf{n}}$ in a power series $f(\mathbf{x})$.

Theorem 16 (multivariate Lagrange inversion formula [1]). *Let g_1, \dots, g_{d+1} , be power series in d variables with non-zero constant terms. There exists a unique tuple (f_1, \dots, f_d) of power series in x_1, \dots, x_d satisfying*

$$(14) \quad f_t(x_1, \dots, x_d) = x_t g_t(f_1, \dots, f_d)$$

for all $t \in [d]$. Moreover, for any tuple $\mathbf{n} = (n_1, \dots, n_d)$ of positive integers

$$[\mathbf{x}^{\mathbf{n}}]g_{d+1}(f_1, \dots, f_d) = [\mathbf{x}^{\mathbf{n}}] \left(\prod_{t \in [d]} \frac{x_t}{n_t} \right) \left(\sum_{A \in \text{Cayley}(d+1, d+1)} \prod_{t \in [d+1]} \left(\prod_{(s, t) \in A} \frac{\partial}{\partial x_s} \right) g_t(\mathbf{x}) \right),$$

where $\text{Cayley}(d+1, d+1)$ is the set of unitype Cayley trees with vertex set $[d+1]$ and root vertex $d+1$, considered as oriented toward their root vertex.

Observe that the sum over $\text{Cayley}(d+1, d+1)$ in Theorem 16 could be written as a determinant of differential operators using the matrix-tree theorem. In [1] it was proved that Theorem 16 is equivalent to some more traditional forms of the multivariate Lagrange inversion formula. This formulation was actually already implicit in [10]. A combinatorial proof of Theorem 16 was first given in [3]. Combinatorial proofs of other forms of the multivariate Lagrange inversion formula were given in [5, 7].

We will now prove Theorem 16 starting from Proposition 11. The existence and uniqueness of power series f_1, \dots, f_d satisfying (14) is clear since their coefficients can be determined inductively from these equations. We now give an interpretation of these series as generating functions of trees (which is equivalent to what is done for instance in [10, Section 1]).

Lemma 17. *For any tuple $\mathbf{n} = (n_1, \dots, n_d)$ of non-negative integers, and for all $\rho \in [d]$, the series f_ρ defined by (14) satisfies*

$$[\mathbf{x}^{\mathbf{n}}]f_\rho = \left(\prod_{t \in [d]} \frac{1}{n_t!} \right) \left(\sum_{T \in \mathcal{T}_\rho(\mathbf{n})} g_{\text{in}(T)} \prod_{s, t \in [d], i \in [n_t]} \text{in}_s(t, i)! \right),$$

where $\mathcal{T}_\rho(\mathbf{n})$ is the set of multitype Cayley trees defined in Section 2, and

$$g_{\text{in}(T)} = \prod_{t \in [d], i \in [n_t]} [y_1^{\text{in}_1(t, i)} \dots y_d^{\text{in}_d(t, i)}] g_t(y_1, \dots, y_d).$$

Proof. For all $\rho \in [d]$, we define the power series \tilde{f}_ρ in the variables x_1, \dots, x_d by

$$\tilde{f}_\rho := \sum_{n_1, \dots, n_d \geq 0} \left(\prod_{t \in [d]} \frac{x_t^{n_t}}{n_t!} \right) \left(\sum_{T \in \mathcal{T}_\rho(\mathbf{n})} g_{\text{in}(T)} \prod_{s, t \in [d], i \in [n_t]} \text{in}_s(t, i)! \right).$$

We want to prove that $\tilde{f}_\rho = f_\rho$. Let $\mathcal{U}_\rho(\mathbf{n})$ be the set of rooted trees obtained by taking a tree T in $\mathcal{T}_\rho(\mathbf{n})$, unlabeled its vertices, and for each vertex v and each type $s \in [d]$ assigning a total order to the children of v of type s . This induces a $\left(\prod_{t \in [d]} n_t! \right)$ -to- $\left(\prod_{s, t \in [d], i \in [n_t]} \text{in}_s(t, i)! \right)$ correspondence between the sets $\mathcal{T}_\rho(\mathbf{n})$ and $\mathcal{U}_\rho(\mathbf{n})$. Thus,

$$(15) \quad \tilde{f}_\rho = \sum_{n_1, \dots, n_d \geq 0} \mathbf{x}^{\mathbf{n}} \sum_{T \in \mathcal{U}_\rho(\mathbf{n})} g_{\text{in}(T)} = \sum_{T \in \mathcal{U}_\rho} \prod_{v \text{ vertex of } T} \omega(v),$$

where $\mathcal{U}_\rho = \bigcup_{n_1, \dots, n_d \geq 0} \mathcal{U}_\rho(\mathbf{n})$, and for a vertex v of type t ,

$$\omega(v) := x_t \cdot [y_1^{\text{in}_1(v)} \dots y_d^{\text{in}_d(v)}] g_t(y_1, \dots, y_d).$$

We now consider the classical decomposition of trees obtained by deleting the root vertex (see for instance [6]). Let $\mathcal{U}_{\rho, (c_1, \dots, c_d)}$ be the set of trees in \mathcal{U}_ρ such that the root vertex has c_s children of type s for all $s \in [d]$. By deleting the root vertex of the trees in $\mathcal{U}_{\rho, (c_1, \dots, c_d)}$, one gets a bijection between $\mathcal{U}_{\rho, (c_1, \dots, c_d)}$ and the Cartesian product

$\mathcal{U}_1^{c_1} \times \cdots \times \mathcal{U}_d^{c_d}$. This gives the following generating function equation

$$\sum_{T \in \mathcal{U}_{\rho, (c_1, \dots, c_d)}} \prod_{v \text{ vertex of } T} \omega(v) = x_{\rho} \cdot \left([y_1^{c_1} \cdots y_d^{c_d}] g_{\rho}(y_1, \dots, y_d) \right) \cdot \tilde{f}_1^{c_1} \cdots \tilde{f}_d^{c_d}.$$

Thus (15) becomes

$$\tilde{f}_{\rho} = x_{\rho} \sum_{c_1, \dots, c_d \geq 0} \left([y_1^{c_1} \cdots y_d^{c_d}] g_{\rho}(y_1, \dots, y_d) \right) \tilde{f}_1^{c_1} \cdots \tilde{f}_d^{c_d} = x_{\rho} g_{\rho}(\tilde{f}_1, \dots, \tilde{f}_d).$$

Since $\tilde{f}_1, \dots, \tilde{f}_d$ are power series satisfying (14), they are equal to f_1, \dots, f_d respectively. \square

We define $f_{d+1} = x_{d+1} g_{d+1}(f_1, \dots, f_d)$, $\mathbf{x}' = (x_1, \dots, x_{d+1})$, $n_{d+1} = 1$, and $\mathbf{n}' = (n_1, \dots, n_{d+1})$. With this notation, Equation (14) is valid for all $t \in [d+1]$, and $[\mathbf{x}^{\mathbf{n}}] g_{d+1}(f_1, \dots, f_d) = [\mathbf{x}'^{\mathbf{n}'}] f_{d+1}$. Hence applying Lemma 17 gives

$$(16) \quad [\mathbf{x}^{\mathbf{n}}] g_{d+1}(f_1, \dots, f_d) = \left(\prod_{t \in [d]} \frac{1}{n_t!} \right) \left(\sum_{T \in \mathcal{T}_{d+1}(\mathbf{n}')} g_{\text{in}(T)} \prod_{s, t \in [d+1], i \in [n_t]} \text{in}_s(t, i)! \right).$$

We now partition the set of trees $\mathcal{T}_{d+1}(\mathbf{n}')$ according to the number $m_{s,t}$ of edges of type (s, t) . Let $M(\mathbf{n})$ be the set of tuples of non-negative integers $(m_{s,t})_{s,t \in [d+1]}$ such that $\sum_{t \in [d+1]} m_{s,t} = n_s$ for all $s \in [d]$, and $m_{d+1,t} = 0$ for all $t \in [d+1]$. For a tuple $\mathbf{m} \in M(\mathbf{n})$, let $\Gamma(\mathbf{m})$ be the set of tuples $(\gamma_{s,t,i})_{s,t \in [d+1], i \in [n_t]}$ of non-negative integers such that $\sum_{i \in [n_t]} \gamma_{s,t,i} = m_{s,t}$ for all $s, t \in [d+1]$. Since $\bigcup_{\mathbf{m} \in M(\mathbf{n})} \Gamma(\mathbf{m})$ is the set of all possible indegree vectors for the trees in $\mathcal{T}_{d+1}(\mathbf{n}')$, we can rewrite the right-hand side of (16) as

$$\text{RHS} = \left(\prod_{t \in [d]} \frac{1}{n_t!} \right) \left(\sum_{\mathbf{m} \in M(\mathbf{n})} \sum_{\gamma \in \Gamma(\mathbf{m})} |\mathcal{T}_{d+1, \gamma}| g_{\gamma} \prod_{s, t \in [d+1], i \in [n_t]} \gamma_{s,t,i}! \right),$$

where

$$g_{\gamma} = \prod_{t \in [d+1], i \in [n_t]} [x_1^{\gamma_{1,t,i}} \cdots x_d^{\gamma_{d,t,i}}] g_t(\mathbf{x}).$$

Using Proposition 11 then gives

$$(17) \quad \text{RHS} = \left(\prod_{t \in [d]} \frac{1}{n_t!} \right) \left(\sum_{\mathbf{m} \in M(\mathbf{n})} \sum_{\gamma \in \Gamma(\mathbf{m})} g_{\gamma} \sum_{A \in \text{Cayley}(d+1, d+1)} \prod_{(s,t) \in A} m_{s,t} \right).$$

Now observe that by definition of $\Gamma(\mathbf{m})$,

$$\sum_{\gamma \in \Gamma(\mathbf{m})} g_{\gamma} = \prod_{t \in [d+1]} [x_1^{m_{1,t}} \cdots x_d^{m_{d,t}}] g_t(\mathbf{x})^{n_t}.$$

Hence for any Cayley tree $A \in \text{Cayley}(d+1, d+1)$,

$$\sum_{\gamma \in \Gamma(\mathbf{m})} g_{\gamma} \prod_{(s,t) \in A} m_{s,t} = \prod_{t \in [d+1]} [x_1^{m_{1,t}} \cdots x_d^{m_{d,t}}] \left(\prod_{(s,t) \in A} x_s \frac{\partial}{\partial x_s} \right) g_t(\mathbf{x})^{n_t},$$

and by definition of $M(\mathbf{n})$,

$$\sum_{\mathbf{m} \in M(\mathbf{n})} \sum_{\gamma \in \Gamma(\mathbf{m})} g_\gamma \prod_{(s,t) \in A} m_{s,t} = [\mathbf{x}^{\mathbf{n}}] \prod_{t \in [d+1]} \left(\prod_{(s,t) \in A} x_s \frac{\partial}{\partial x_s} \right) g_t(\mathbf{x})^{n_t}.$$

Moreover, $\prod_{(s,t) \in A} x_s = \prod_{t \in [d]} x_t$. Thus (17) gives

$$\text{RHS} = [\mathbf{x}^{\mathbf{n}}] \left(\prod_{t \in [d]} \frac{x_t}{n_t} \right) \left(\sum_{A \in \text{Cayley}(d+1, d+1)} \prod_{t \in [d+1]} \left(\prod_{(s,t) \in A} \frac{\partial}{\partial x_s} \right) g_t(\mathbf{x})^{n_t} \right).$$

This completes the proof of Theorem 16.

5. ENUMERATION OF PLANE TREES AND CACTI

In this section we investigate the enumeration of plane trees and cacti using the same philosophy behind the proof of Theorem 11. Recall that a *rooted plane tree* is a rooted tree in which the children of each vertex are ordered. A *vertex-labeled* rooted plane tree is a rooted plane tree in which the n vertices receive distinct labels in $[n]$; equivalently it is a rooted Cayley tree in which the children of each vertex are ordered.

As mentioned in the introduction, the problem of counting rooted plane trees is equivalent to the problem of counting rooted Cayley trees. Let us illustrate our point by enumerating rooted (untype) plane trees having N_i vertices with i children. By the case $d = 1$ of (10), for any tuple of non-negative integers $\gamma = (\gamma_1, \dots, \gamma_n)$ summing to $n - 1$, the number of rooted Cayley trees with n vertices in which vertex i has γ_i children for all $i \in [n]$ is

$$\binom{n-1}{\gamma_1, \gamma_2, \dots, \gamma_n}.$$

This means that the set \mathcal{P}_γ of vertex-labeled rooted plane trees in which vertex i has γ_i children has cardinality

$$(18) \quad |\mathcal{P}_\gamma| = (n-1)!.$$

Let $\mathbf{N} = (N_0, N_1, \dots)$ be a sequence of non-negative integers summing to n , and let $\mathcal{L}^{\mathbf{N}}$ be the set of vertex-labeled rooted plane trees having N_i vertices with i children for all $i \geq 0$. This set is empty unless $\sum_{i \geq 0} N_i = 1 + \sum_{i \geq 0} i N_i$. Moreover, in this case from (18) we immediately get

$$|\mathcal{L}^{\mathbf{N}}| = (n-1)! \binom{n}{N_0, N_1, N_2, \dots}.$$

Now, since a rooted plane tree has no non-trivial automorphisms, there are $n!$ ways of labeling its n vertices. Thus the number of *unlabeled* rooted plane trees having N_i vertices with i children for all i is

$$(19) \quad |\mathcal{P}^{\mathbf{N}}| = \frac{1}{n} \binom{n}{N_0, N_1, N_2, \dots},$$

if $\sum_{i \geq 0} N_i = 1 + \sum_{i \geq 0} i N_i$ and 0 otherwise. Classically, this result is obtained using Łukasiewicz words and the Cycle Lemma (see e.g. [18, Section 5.3]).

We could also have proved (19) directly, that is, without using (10). Indeed, by a reasoning analogous to Lemma 4 one easily shows that $|\mathcal{P}_\gamma| = |\mathcal{P}_{\gamma'}|$ for any tuples $\gamma = (\gamma_1, \dots, \gamma_n)$ and $\gamma' = (\gamma'_1, \dots, \gamma'_n)$ of non-negative integers summing to $n - 1$. This, in turn, immediately gives (18). We will not dwell further into this parallel approach for counting plane trees.

In the rest of this section we deal with the enumeration of planar cacti. We give precise definitions below, but roughly speaking, a (planar) d -cactus is a tree-like structure in which edges are replaced by d -gons (polygons with d sides). Examples are shown in Figure 4. We shall enumerate planar d -cacti according to the degree distribution of their vertices. This enumeration is equivalent to the computation of certain connection coefficients in the symmetric group (corresponding to the “minimal” factorizations of the long cycle into d permutations, see [9] or [14, Section 1.3]). The results we obtain are not new: they have been obtained by the Lagrange inversion formula in [9, 2], representation theory [11], and even combinatorial methods [3]. Our contribution is to give a shorter proof, which takes advantage of the symmetries in the counting formulas for d -cacti.

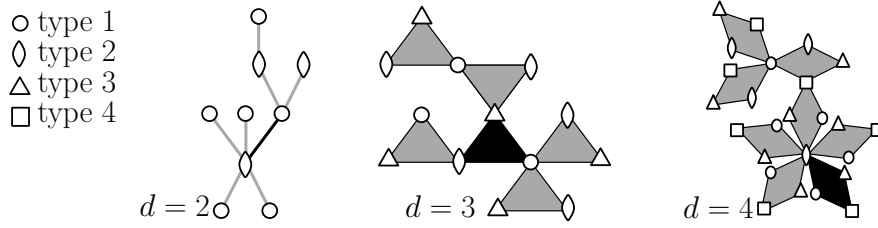


FIGURE 4. Examples of rooted planar d -cacti for $d = 2, 3, 4$. In all figures, the root d -gon is colored in black.

We fix an integer $d \geq 2$. A d -graph is a connected simple graph with d types of vertices $1, 2, \dots, d$, such that every 2-connected component is a cycle of length d with vertices of type $1, 2, \dots, d$ appearing in this order cyclically. Each such cycle is called a d -gon. A *plane embedding* of a d -graph is a drawing of d -graph in the plane without edge-crossings in such a way that every edge is incident to the unbounded region, and around each d -gon the vertices of type $1, 2, \dots, d$ appear in clockwise order. A (planar) d -cacti is a plane-embedding of a d -graph considered up to continuous deformation (this is equivalent to fixing for each vertex v the cyclic order of its incident d -gons). The *size* of a d -cactus is its number of d -gons. Note that a d -cactus of size n has dn edges and, by the Euler relation, $(d - 1)n + 1$ vertices. The *degree* of a vertex is the number of incident d -gons.

A d -cactus is *rooted* if one of the d -gons is distinguished as the *root d -gon*. Let $\mathbf{n} = (n_1, \dots, n_d)$ be a tuple of positive integers. A d -cactus with n_t vertices of type t for all $t \in [d]$ is said to be *vertex-labeled* if the vertices of type t have distinct labels in $[n_t]$ for all $t \in [d]$. The labeled vertices are denoted by (t, i) where $t \in [d]$ is the type and $i \in [n_t]$ is the label. We denote by $\mathcal{C}(\mathbf{n})$ the set of rooted vertex-labeled planar d -cacti with n_t vertices of type t . As mentioned earlier, the set $\mathcal{C}(\mathbf{n})$ is empty unless $n := \left(\sum_{t \in [d]} n_t - 1 \right) / (d - 1)$ is an integer which is the size of the cacti. Moreover we

must impose $n_t \leq n$ for all $t \in [d]$ and we now suppose that these conditions hold. For a tuple of integers $\mathbf{N} = (N_{t,j})_{t \in [d], j > 0}$ we denote by $\mathcal{C}^{\mathbf{N}}(\mathbf{n})$ the set of d -cacti in $\mathcal{C}(\mathbf{n})$ having $N_{t,j}$ vertices of type t and degree j for all $t \in [d], j > 0$. Of course this set is empty unless

$$\sum_{j>0} N_{t,j} = n_t, \quad \text{and} \quad \sum_{j>0} j N_{t,j} = n$$

for all $t \in [d]$. In [9] Goulden and Jackson proved that under these conditions

$$(20) \quad |\mathcal{C}^{\mathbf{N}}(\mathbf{n})| = n^{d-1} \frac{\prod_{t \in [d]} n_t! (n_t - 1)!}{\prod_{t \in [d], j > 0} N_{t,j}!}.$$

Actually the result in [9] is stated in terms of rooted d -cacti with *unlabeled* vertices, accounting for a factor $\prod_{t \in [d]} n_t!$ between (20) and [9, Theorem 3.2].

We will now give a proof of (20) using a philosophy similar to the one developed in previous sections. For a cactus $C \in \mathcal{C}(\mathbf{n})$, we denote by (t, i) the vertex of type t labeled i . For a tuple $\gamma = (\gamma_{t,i})_{t \in [d], i \in [n_t]}$ of integers we say that C has *degree vector* γ if the vertex (t, i) has degree $\gamma_{t,i}$ for all $t \in [d], i \in [n_t]$. We denote by \mathcal{C}_γ the set of d -cacti in $\mathcal{C}(\mathbf{n})$ having degree vector γ . Of course this set is empty unless $\sum_{i \in [n_t]} \gamma_{t,i} = n$ for all $t \in [d]$ and $\gamma_{t,i} > 0$ for all $t \in [d], i > 0$. An equivalent way of stating (20) is as follows.

Theorem 18. *Let n_1, \dots, n_d be positive integers such that $n := (\sum_{t \in [d]} n_t - 1) / (d - 1)$ is an integer, and $n_t \leq n$ for all $t \in [d]$. Let $\gamma = (\gamma_{t,i})_{t \in [d], i \in [n_t]}$ be a tuple of positive integers such that $\sum_{i \in [n_t]} \gamma_{t,i} = n$ for all $t \in [d]$. Then the number of rooted vertex-labeled d -cacti with degree vector γ is*

$$(21) \quad |\mathcal{C}_\gamma| = n^{d-1} \prod_{t \in [d]} (n_t - 1)!.$$

Corollary 19. *The number of rooted vertex-labeled d -cacti of size n with n_t vertices of type t for all $t \in [d]$ is*

$$|\mathcal{C}(\mathbf{n})| = n^{d-1} \prod_{t \in [d]} \binom{n-1}{n_t-1} (n_t - 1)!,$$

if $(\sum_{t \in [d]} n_t - 1) / (d - 1) = n$, and 0 otherwise.

Proof of Corollary 19. The right-hand side of (21) does not depend on γ . And since $|\mathcal{C}(\mathbf{n})| = \sum_\gamma |\mathcal{C}_\gamma|$, then it suffices to observe that for all $t \in [d]$ there are $\binom{n-1}{n_t-1}$ tuples of positive integers such that $\sum_{i \in [n_t]} \gamma_{t,i} = n$. \square

The rest of this section is devoted to the proof of Theorem 18. We start with an analogue of Lemma 4.

Lemma 20. *Let $\mathbf{n} = (n_1, \dots, n_d)$ be a tuple of positive integers such that $n := (\sum_{t \in [d]} n_t - 1) / (d - 1)$ is an integer. Let $\gamma = (\gamma_{t,i})_{t \in [d], i \in [n_t]}$ and $\gamma' = (\gamma'_{t,i})_{t \in [d], i \in [n_t]}$ be tuples of positive integers such that $\sum_{i \in [n_t]} \gamma_{t,i} = \sum_{i \in [n_t]} \gamma'_{t,i} = n$ for all $t \in [d]$. Then the number of d -cacti in $\mathcal{C}(\mathbf{n})$ having degree vector γ or γ' is the same: $|\mathcal{C}_\gamma| = |\mathcal{C}_{\gamma'}|$.*

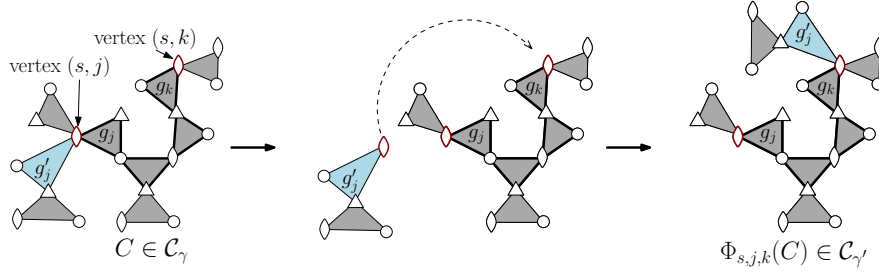


FIGURE 5. The bijection $\Phi_{s,j,k}$. The chain of d -gons between vertices (s,j) and (s,k) is drawn in thick lines.

Proof. It suffices to prove $|\mathcal{C}_\gamma| = |\mathcal{C}_{\gamma'}|$ when the tuples γ and γ' only differ on two coordinates. Let $s \in [d]$, $j, k \in [n_s]$ such that $\gamma_{s,j} > 1$, let γ' be defined by $\gamma'_{s,j} = \gamma_{s,j} - 1$, $\gamma'_{s,k} = \gamma_{s,k} + 1$ and $\gamma'_{t,i} = \gamma_{t,i}$ for all $(t,i) \notin \{(s,j), (s,k)\}$. We need to prove $|\mathcal{C}_\gamma| = |\mathcal{C}_{\gamma'}|$ and we proceed to exhibit a bijection $\Phi_{s,j,k}$ between \mathcal{C}_γ and $\mathcal{C}_{\gamma'}$. The bijection $\Phi_{s,j,k}$ is illustrated in Figure 5 (the root d -gon plays no role in this construction and is not indicated).

Let C be a cactus in \mathcal{C}_γ . Let P be the *chain of d -gons* between the vertices (s,j) and (s,k) , that is, the set of d -gons containing the edges of any simple path between the vertices (s,j) and (s,k) (this set of d -gons is independent of the simple path considered); see Figure 5. Let g_j (resp. g_k) be the d -gon in P incident to the vertex (s,j) (resp. (s,k)). Let g'_j be the d -gon incident to the vertex (s,j) following g_j clockwise around (s,j) . Observe that $g'_j \neq g_j$ since the vertex (s,j) has degree $\gamma_{s,j} > 1$. We define $\Phi_{s,j,k}(C)$ as the cactus obtained from C by ungluing the d -gon g'_j from the vertex (s,j) and gluing it to (s,k) in the corner following the d -gon g_j clockwise around (s,k) ; see Figure 5. It is clear that $\Phi_{s,j,k}(C)$ is a cactus in $\mathcal{C}_{\gamma'}$. It is equally clear that $\Phi_{s,j,k}$ is a bijection between \mathcal{C}_γ and $\mathcal{C}_{\gamma'}$, since the inverse mapping is $\Phi_{s,j,k}^{-1} = \Phi_{s,k,j}$. This completes the proof. \square

Given Lemma 20 it is sufficient to prove (21) for the particular tuple $\gamma^*(\mathbf{n}) = (\gamma_{t,i})_{t \in [d], i \in [n_t]}$ defined by $\gamma_{t,1} = n - n_t + 1$ and $\gamma_{t,i} = 1$ for all t in $[d]$ and all $i > 1$ in $[n_t]$. Cacti in $\mathcal{C}_{\gamma^*(\mathbf{n})}$ are shown in Figure 7. We will now invoke a second symmetry in order to enumerate $\mathcal{C}_{\gamma^*(\mathbf{n})}$.

Lemma 21. *Let $\mathbf{n} = (n_1, \dots, n_d)$ be a tuple of positive integers. Let $r, s \in [d]$ be such that $n_r > 1$, and $n_s < n$ and let $\mathbf{n}' = (n'_1, \dots, n'_d)$ be defined by $n'_r = n_r - 1$, $n'_s = n_s + 1$ and $n'_t = n_t$ for all $t \notin \{r, s\}$. Then*

$$|\mathcal{C}_{\gamma^*(\mathbf{n})}| / (n_r - 1) = |\mathcal{C}_{\gamma^*(\mathbf{n}')}| / (n'_s - 1).$$

Proof. Let C be a cactus in $\mathcal{C}_{\gamma^*(\mathbf{n})}$. Let P be the *chain of d -gons* between vertices $(r,1)$ and $(s,1)$ (the set of d -gons containing the edges of any simple path between these vertices). Let g_s be the unique d -gon in P incident to the vertex $(s,1)$ and let g'_s be the d -gon incident to $(s,1)$ following g_s clockwise around $(s,1)$. Observe that $g'_s \neq g_s$ since the vertex $(s,1)$ has degree $n - n_s + 1 > 1$. Hence, the vertex v of type r incident to g'_s has a label distinct from 1, and has degree 1. We say that the cactus $C \in \mathcal{C}_{\gamma^*(\mathbf{n})}$ is in

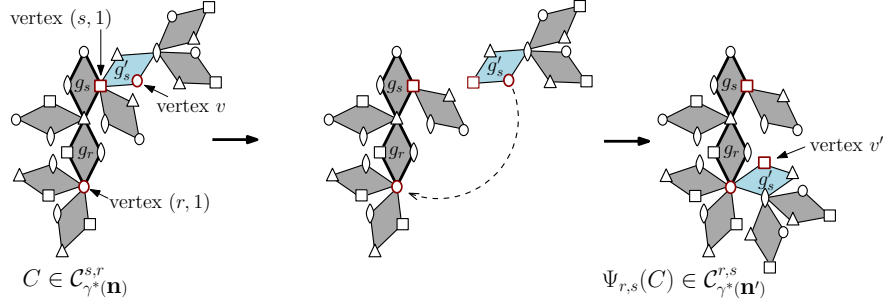


FIGURE 6. The bijection $\Psi_{r,s}$. The chain of d -gons between vertices $(r,1)$ and $(s,1)$ is drawn in thick lines. The d -gon g'_s is indicated in lighter shade. The vertex $v = (r, n_r)$ and $v' = (s, n_s + 1)$.

$\mathcal{C}_{\gamma^*}^{s,r}(\mathbf{n})$ if v has label n_r . It is clear that $|\mathcal{C}_{\gamma^*}^{s,r}(\mathbf{n})| = |\mathcal{C}_{\gamma^*}(\mathbf{n})|/(n_r - 1)$. Hence to prove the lemma it suffices to exhibit a bijection $\Psi_{r,s}$ between $\mathcal{C}_{\gamma^*}^{s,r}(\mathbf{n})$ and $\mathcal{C}_{\gamma^*}^{r,s}(\mathbf{n}')$. This bijection is represented in Figure 6.

Let C be a cactus in $\mathcal{C}_{\gamma^*}^{s,r}(\mathbf{n})$. Let P , g_s , g'_s and v be as above. We also denote by g_r be the unique d -gon in the chain P incident to the vertex $(r,1)$. We then denote by $\Psi_{r,s}(C)$ the cactus obtained from C by ungluing the d -gon g'_s from the vertex $(s,1)$, and gluing it to the vertex $(r,1)$ in the corner following the d -gon g_r around $(r,1)$; see Figure 6. In this process, the vertex v becomes identified with the vertex $(s,1)$ (so that the label n_r of v disappears), while the vertex v' of type s incident to the d -gon g'_s (previously identified with $(s,1)$) takes the label $n_s + 1$. It is clear that $\Psi_{r,s}(C)$ is a cactus in $\mathcal{C}_{\gamma^*}^{r,s}(\mathbf{n}')$. It is equally clear that $\Psi_{r,s}$ is a bijection between $\mathcal{C}_{\gamma^*}^{s,r}(\mathbf{n})$ and $\mathcal{C}_{\gamma^*}^{r,s}(\mathbf{n}')$, since the inverse mapping is $\Psi_{r,s}^{-1} = \Psi_{s,r}$. This completes the proof. \square

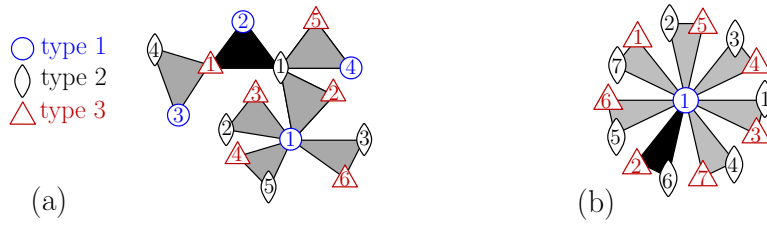


FIGURE 7. (a) A cactus in $\mathcal{C}_{\gamma^*}(\mathbf{n})$, for $\mathbf{n} = (4,4,5)$. (b) A cactus in $\mathcal{C}_{\gamma^*}(\mathbf{n})$, for $\mathbf{n} = (1,7,7)$. In both figures the root d -gon is colored in black and the numbers indicate the labels of vertices.

By Lemma 21, we immediately get that for any tuple $\mathbf{n} = (n_1, \dots, n_d)$ in $[n]^d$ such that $n := \left(\sum_{t \in [d]} n_t - 1\right)/(d-1)$ is an integer,

$$|\mathcal{C}_{\gamma^*}(\mathbf{n})| = \frac{\prod_{t \in [d]} (n_t - 1)!}{(n-1)!^{d-1}} |\mathcal{C}_{\gamma^*}^{*(1,n,n,\dots,n)}|.$$

Moreover the set of cacti $\mathcal{C}_{\gamma^*(1,n,n,\dots,n)}^*$ is very easy to enumerate. Indeed this is the set of cacti where the unique vertex of type 1 has degree n and all the other vertices have degree 1; see Figure 7(b). The only freedom left in such a cactus are the labels of the n vertices of type t for all $t \in \{2, \dots, d\}$. This gives

$$|\mathcal{C}_{\gamma^*(1,n,n,\dots,n)}^*| = n!^{d-1},$$

hence $|\mathcal{C}_{\gamma^*(n)}| = n^{d-1} \prod_{t \in [d]} (n_t - 1)!$. Together with Lemma 20 this completes the proof of Theorem 18.

Acknowledgment: We thank Mireille Bousquet-Mélou for pointing out an error in a previous version of this manuscript, Ira Gessel for mentioning reference [13], and to anonymous referees for valuable suggestions.

REFERENCES

- [1] E.A. Bender and L.B. Richmond. A multivariate Lagrange inversion formula for asymptotic calculations. *Electron. J. Combin.*, 5(R33), 1998.
- [2] M. Bóna, G. Bousquet, M. Labelle, and P. Leroux. Enumeration of m -ary cacti. *Adv. in Appl. Math.*, 24(1):22–56, 2000.
- [3] M. Bousquet, C. Chauve, G. Labelle, and P. Leroux. Two bijective proofs for the arborescent form of the Good-Lagrange formula and some applications to colored rooted trees and cacti. *Theor. Comput. Sci.*, 307(2):277–303, 2003.
- [4] M. Bousquet-Mélou and G. Chapuy. The vertical profile of embedded trees. *Electron. J. Combin.*, 19(P46), 2012.
- [5] R. Ehrenborg and M. Méndez. A bijective proof of infinite variated Good’s inversion. *Adv. in Math.*, 103(2):221–259, 1994.
- [6] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, 2009.
- [7] I.M. Gessel. A combinatorial proof of the multivariable Lagrange inversion formula. *J. Combin. Theory Ser. A*, 45(2), 1987.
- [8] I.J. Good. Generalizations to several variables of Lagrange’s expansion, with applications to stochastic processes. *Proc. Cambridge Philos. Soc.*, 56:367–380, 1960.
- [9] I.P. Goulden and D.M. Jackson. The combinatorial relationship between trees, cacti and certain connection coefficients for the symmetric group. *European J. Combin.*, 13(5):357–365, 1992.
- [10] I.P. Goulden and D.M. Kulkarni. Multivariable Lagrange inversion, Gessel-Viennot cancellation, and the matrix tree theorem. *J. Combin. Theory Ser. A*, 80(2):295–308, 1997.
- [11] D.M. Jackson. Some combinatorial problems associated with products of conjugacy classes of the symmetric group. *J. Combin. Theory Ser. A*, 49(2), 1988.
- [12] A. Joyal. Une théorie combinatoire des séries formelles. *Adv. in Math.*, 42:1–82, 1981.
- [13] D.E. Knuth. Another enumeration of trees. *Canad. J. Math.*, 20:1077–1086, 1968.
- [14] S.K. Lando and A.K. Zvonkin. *Graphs on surfaces and their applications*. Springer-Verlag, 2004.
- [15] J.W. Moon. *Counting Labelled Trees*. Canadian Mathematical Monographs, 1970.
- [16] J. Pitman. Coalescent random forests. *J. Combin. Theory Ser. A*, 85(2):165–193, 1999.
- [17] H. Prüfer. Neuer Beweis eines Satzes über Permutationen. *Arch. Math. Phys.*, 27:742–744, 1918.
- [18] R.P. Stanley. *Enumerative combinatorics, volume 2*. Cambridge University Press, 1999.